·论著·

中国老年人慢性病多病共存模式的研究

潘晔[1]，刘志辉[2]，胡倩倩[3]，王留义[2*]

【摘要】 背景 随着人口老龄化及寿命延长，慢性病的多病共存日益普遍。疾病种类多且病情复杂，为老年人健康管理提出挑战。共病模式作为研究的必要问题，国内研究相对缺乏。**目的** 研究中国老年人常见慢性病的多病共存模式，帮助政策制定者、研究人员和临床医生更好地了解多病共存现状。**方法** 选取中国健康养老追踪调查（CHARLS）2018 年数据中 60 岁及以上被访者，选择人口特征学部分数据及健康状况的 14 种慢病数据。采用关联规则、聚类分析、主成分分析、潜在类别分析 4 种方法对中国老年人的共病模式进行探索，并对比不同方法所得结果。**结果** 最终共纳入 10 800 例被访者数据。4 种方法所得模式有所差异，但存在一致性的共病模式：（1）高血压、糖尿病或血糖升高、血脂异常；（2）慢性肺部疾患和哮喘；（3）关节炎或风湿病、胃部疾病或消化系统疾病；（4）中风、与记忆相关的疾病。**结论** 不同方法得到的一致模式，所包含慢病病因关系关联明显；共病模式存在差异的原因是包含的病因关系复杂，且方法的原理不同。

【关键词】 慢性病；慢性病共病；共病现象；关联规则；系统聚类；潜在类别分析

【中图分类号】 R 36 【文献标识码】 A DOI：10.12114/j.issn.1007-9572.2022.0820

【引用本文】 潘晔，刘志辉，胡倩倩，等．中国老年人慢性病多病共存模式的研究［J］．中国全科医学，2023.［Epub ahead of print］．DOI：10.12114/j.issn.1007-9572.2022.0820.［www.chinagp.net］

PAN Y，LIU Z H，HU Q Q，et al. Patterns of coexistence of multiple chronic conditions among Chinese elderly［J］. Chinese General Practice，2023.［Epub ahead of print］.

Patterns of Coexistence of Multiple Chronic Conditions among Chinese Elderly

PAN Ye[1]，*LIU Zhihui*[2]，*HU Qianqian*[3]，*WANG Liuyi*[2*]

1.*Henan Integrative Medicine Hospital*，*Henan*，*Zhengzhou* 450004，*China*

2.*Department of General Practice*，*Henan Provincial People's Hospital*，*Henan*，*Zhengzhou* 450003，*China*

3.*Henan UniversityPeople's Hospital*，*Zhengzhou* 450003，*China*

**Corresponding author*：*WANG Liuyi*，*Chief physician*，*Email*：*wly*2000@126.com

【Abstract】 **Background** With the aging and longer survival of the population，comorbid chronic diseases is increasingly common. The variety and complexity of diseases poses challenges to the health management of the elderly.There is a relative lack of multimorbidity pattern researches in China，whichare necessary issues for research. **Objective** To investigate the patterns of coexistence of common multiple chronic conditions among the elderly in China，in order to help policymakers，researchers，and clinicians better understand the current status of multimorbidity among Chinese elderly. **Methods** Data on the demographic characteristics and health status of 14 chronic diseases were extracted from the respondents aged 60 years and above in the China Health and Retirement Longitudinal Study（CHARLS）2018，association rules，cluster analysis，principal component analysis，latent class analysis were used to explore the multimorbidity patterns of Chinese elderly，and the results of different methods were compared. **Results** The data from a total of 10 800 respondents were eventually included，there were differences among the patterns obtained by four methods.However，the consistent multimorbidity patterns were identified：（i）hypertension，diabetes or elevated blood glucose，dyslipidemia；（ii）chronic lung disease and asthma；（iii）arthritis or rheumatism，stomach diseases or digestive diseases；（iv）stroke，memory-related diseases. **Conclusion** The consistent patterns obtained by different methods contain chronic diseases with significant relationships of etiologies.The reasons of differences in results are complex etiologic relationships and different method principles.

【Key words】 Chronic disease；Multiple chronic conditions；Comorbidity；Association rules；Clustering analysis；Latent class analysis

---

基金项目：河南省中医药科学研究专项课题（2023ZYZD13）；河南省中医药文化与管理研究项目（TCM2023005）

1.450004 河南省郑州市，河南省中西医结合医院 2.450003 河南省郑州市，河南省人民医院全科医学科 3.450003 河南省郑州市，河南大学人民医院

*通信作者：王留义，主任医师；E-mail：wly2000@126.com

本文数字出版日期：2023-05-26

WHO 对于多病共存（multimorbidity）的定义为：一个人同时患有两种或多种疾病[1]。城镇化及老龄化使老年人疾病谱发生了显著变化，越来越多的人同时患有多种疾病。多病共存涉及疾病的病因关系多样，疾病类型没有限定，并不明确哪一种疾病为索引疾病[2]。国内外多病共存模式研究常用的方法有：多病个数统计[3]、聚类分析[4]、因子分析[5]及相关分析[6]。研究数据来源为：调查数据、临床数据或医疗数据，涉及疾病有些是常见疾病，有些是罕见疾病。孙志佳等[7]选取 2 次中国慢病前瞻性研究数据，采用个数统计及系统聚类对 10 个地区被访者的共病模式进行分析；刘贝贝等[8]选取 CHARLS 数据采用关联规则探索常见慢性疾病的共病组合模式；OLAYA 等[9]采用潜在类别分析区分“Edad con Salud”调查被访者共病模式并进行预后分析，涉及疾病包含常见慢病及白内障。既往研究选取数据、疾病及使用方法不同，得到的共存模式研究结果各异，可信度较低，很少有对同一数据采取多种方法，并对比结果来确保研究成果的可靠性。

本研究旨在采用多种方法探索中国老年人常见慢病的多病共存模式，并分析不同方法所得结果的差异性与一致性。选取中国健康与养老追踪调查（China Health and Retirement Longitudinal Study，CHARLS）2018 年的数据，应用 4 种方法探索多病模式，避免了不同数据导致结果无法比较的问题，通过比较不同方法所得结果的一致性，从而得出稳定的多病模式。

1 资料与方法

1.1 数据来源 CHARLS 旨在收集代表中国 45 岁及以上中老年人家庭和个人的高质量微观数据，用以分析我国人口老龄化问题，推动老龄化问题的跨学科研究[10-12]。选择经过数据清洗的 Harmonized_CHARLS_D 版本，从中筛选 2018 年的数据进行分析，选取人口特征学部分数据及健康状况的 14 种慢病数据。

1.2 数据处理 由于 Harmonized_CHARLS_D 版本经过数据清洗，质量较高。为了保证研究数据与总体的一致性，仅筛选年龄 60 岁及以上的被访者，不删除缺失值，最终得到 10 800 份被访者数据。14 种慢病数据的缺失值较少，通过众数填补。人口特征学数据未参与模式分析故不做处理，在描述时以缺失值展示。

1.3 数据定义方法 CHARLS 数据选取的 14 种慢病为：高血压，糖尿病或血糖升高（包括糖耐量异常和空腹血糖升高），癌症等恶性肿瘤（不包括轻度皮肤癌），慢性肺部疾患如慢性支气管炎或肺气肿、肺心病（不包括肿瘤或癌，后简称慢性肺部疾患），心脏病（如心肌梗死、冠心病、心绞痛、充血性心力衰竭和其他心脏疾病），中风，情感及精神方面问题，关节炎或风湿病，血脂异常（高血脂或低血脂），肝脏疾病（除脂肪肝、肿瘤或癌外），肾脏疾病（不包括肿瘤或癌），胃部疾病或消化系统疾病（不包括肿瘤或癌），哮喘，与记忆相关的疾病（如老年痴呆症、脑萎缩、帕金森症）。

人口特征学数据包括：性别（男、女），年龄，家庭居住地类型（城市、农村），被访者户口类型（城市户口、农村户口），受教育程度（小学以下、小学及初中学历、高中及职业培训、大专及以上）及婚姻状况（已婚、分居或离婚、丧偶、未婚）。

1.4 研究方法 数据处理及所用方法通过 R 语言编程实现，运用多个 R 语言包完成 4 种分析。数据的一般情况包含被访者人群的概况统计，14 种慢病的发病率、共病率、平均共病个数及方差，以此了解被访者人群的特征及 14 种慢病的发病情况。

1.4.1 关联规则 采用关联规则发现常见两病及三病组合，支持度为疾病组合的观察发病率，提升度为独立性判断指标，在两病组合中与观察发病率 / 期望发病率相同。设置关联规则前项集与后项集的个数、支持度最小值可以得到观察发病率较高的组合。计算组合预期发病率并结合提升度，判断疾病的独立性。

1.4.2 聚类分析 聚类分析根据某些数量特征将观察对象分类，根据聚类对象的不同，可以分为个案聚类和变量聚类。反映个案或者变量间性质远近的统计量称为距离，有多种计算方法。将被访者的 14 种慢病数据处理为 01 形式，即被访者患有此疾病时为 1，未患有此疾病时为 0。处理完成的 14 种慢病数据可以看成一个 n*p 的 01 矩阵，其中 n 是被访者的个数（10 800），p 为慢病种数 14。被访者的聚类是个案聚类，依据 14 种慢病的发病情况计算得到的 n*n 距离矩阵进行分类。对 14 种慢病的系统聚类通过慢病在被访者中的共现情况，得到一个 p*p 距离矩阵并由此分类。本研究对被访者采用 k-medoids 算法，对 14 种慢病采用系统聚类以区分多病共存模式。

01 数据距离的选择取决于正匹配（同为 1）和负匹配（同为 0）的相对权重，而相对权重又取决于正匹配和负匹配的相对重要性。本研究采用 01 数据常用的 Jaccard 距离和 Yule's Q 距离，Yule's Q 是一种能体现两种慢病负匹配及未匹配时信息的相似度计算方法，较好地反映 01 数据所有相关信息[13]，Jaccard 距离仅突出正匹配[14]。共病模式中侧重于慢病正匹配，即共同发病，并且被访者中存在健康者（全 0 数据），Yule's Q 在 1 个数据为全 0 时无法计算。对比两个距离计算方法的效果，并考虑其临床意义后，最终采用 Yule's Q 距离进行慢病的系统聚类分析；采用 Jaccard 距离进行被访者的 k-medoids 聚类。系统聚类的分类个数通过谱系图结合临床意义确定，聚类方法为类平均法；被访者的 k-medoids 聚类个数通过 factoextra 包中的 fviz_nbclust 函

数确定。

1.4.3 主成分分析 每个变量在不同程度上反映了某些信息并存在有一定的相关性。变量太多会增加计算量及分析问题的复杂性，主成分分析通过恰当的数学变换，使新变量为原变量的线性组合，并选取少数几个在变差总信息量中比例较大的主成分来分析事物。本研究采用最大方差旋转的主成分分析方法，使用方差贡献度确定主成分个数并将 14 个多病化简为少数疾病组合，以载荷系数绝对值 >0.36 选择主成分包含的慢病。

1.4.4 潜在类别分析（latent class analysis，LCA） LCA 根据被访者在 14 种慢病的患病情况对其分类。与聚类分析类似，LCA 通过个体间相似性区分类别，但不定义类别之间的距离，也不选择分类方法，根据个体所有变量的观测变量概率进行分类。遍历所有的分类，比较指标得到最优分类个数。常用于判别最优模型的指标是：Likelihood-ratio，G2 statistic，Bayesian Information Criterion（BIC）等，采用 R 语言的 poLCA 包进行 LCA 分析，最终类别判定指标为 BIC 值。

2 结果

2.1 被访者人群特征 最终得到 10 800 例 60 岁及以上的被访者，被访者中女性稍多（51.3%），居住在农村的被访者较多（60.3%），农村户口被访者较高（74.0%）。被访者受教育程度普遍较低，90.7% 的被访者未接受高中及以上教育，由于 CHARLS 访谈以家庭为单位进行，已婚被访者较多（77.9%），这与国内人口总体特征一致[15]。86.2% 的被访者患有 1 种及以上的慢病，49.0% 的被访者患有 2~4 种慢病（表 1）。

14 种慢病共病率均高于 85%（表 2），发病率最高的是高血压（47.14%），平均共病个数较多的是情感及精神方面问题（4.16）及与记忆相关的疾病（4.17）。中风、肝脏疾病及肾脏疾病的发病率较低，但平均共病个数较高。

表 1 研究对象的慢病种数及基本特征

Table 1 Number and basic characteristics of chronic diseases of research subjects

组别	例数	多病种数	无慢病	1 种慢病	2 种慢病	3 种慢病	4 种慢病	2018 年龄		家庭居住地类型	
								年龄（$\bar{x} \pm s$，岁）	年龄〔M（P_{25}，P_{75}），岁〕	城市	农村
性别											
男	5 261	826（15.7）	1 181（22.4）	1 082（20.6）	868（16.5）	564（10.7）	740（14.1）	69.1 ± 7.0	68.0（60.0，118.0）	2 056（39.1）	3 205（60.9）
女	5 539	660（11.9）	1 077（19.4）	1 164（21.0）	960（17.3）	660（11.9）	1 018（18.4）	69.2 ± 7.2	68.0（60.0，108.0）	2 234（40.3）	3 305（59.7）
总体	10 800	1 486（13.8）	2 258（20.9）	2 246（20.8）	1 828（16.9）	1 224（11.3）	1 758（16.3）	69.2 ± 7.2	68.0（60.0，118.0）	4 290（39.7）	6 510（60.3）

组别	被访者户口类型			受教育程度			婚姻状况			
	城市户口	农村户口	缺失值	小学以下、小学及初中	高中及职业培训	大专及以上	已婚	分居或离婚	丧偶	未婚
性别										
男	1 401（26.6）	3 755（71.4）	105（2.0）	4 565（86.8）	578（11.0）	118（2.2）	4 530（86.1）	73（1.4）	592（11.3）	66（1.3）
女	1 179（21.3）	4 234（76.4）	126（2.3）	5 234（94.5）	263（4.7）	42（0.8）	3 883（70.1）	54（1.0）	1 597（28.8）	5（0.1）
总体	2 580（23.9）	7 989（74.0）	231（2.1）	9 799（90.7）	841（7.8）	160（1.5）	8 413（77.9）	127（1.2）	2 189（20.3）	71（0.7）

2.2 关联规则分析 表 3 展示了支持度 >10% 的所有两病组合，表 4 展示了支持度 >5% 的三病组合。仅从支持度考虑，关联规则得到的共病组合仅是共同发生，例如：关节炎或风湿病与高血压在老年人群中共同发病率较高，但病因相对独立，故观察发病率 / 期望发病率接近于 1。血脂异常与心脏病共同发病率高，结合观察发病率与期望发病率的发现关联，两者病因有关。

三病组合选取观察发病率与期望发病率的比值及提升度来判断疾病间的联系。一：心脏病，血脂异常及高血压病；二：糖尿病或血糖升高，血脂异常及高血压病；两个病组的观察发病率 / 期望发病率高，提升度较高，在病因学关系及致病因素上存在关联。

表 2 慢病发病率及共病率

Table 2 Incidence and co-morbidity of chronic diseases

慢病	发病人数（例）	共病人数（例）	平均共病种数（$\bar{x} \pm s$）
高血压	5 091（47.14）	4 458（87.57）	2.56 ± 1.92
糖尿病或血糖升高	1 656（15.33）	1 578（95.29）	3.41 ± 2.06
癌症等恶性肿瘤	279（2.58）	249（89.25）	3.09 ± 2.44
慢性肺部疾患	2 121（19.64）	1 983（93.49）	3.34 ± 2.09
心脏病	2 684（24.85）	2 580（96.13）	3.46 ± 1.98
中风	1 093（10.12）	1 047（95.79）	3.51 ± 2.07
情感及精神方面问题	398（3.69）	377（94.72）	4.16 ± 2.56
关节炎或风湿病	4 808（44.52）	4 159（86.50）	2.53 ± 1.96
血脂异常	2 644（24.48）	2 531（95.73）	3.37 ± 1.97
肝脏疾病	779（7.21）	745（95.64）	3.87 ± 2.29
肾脏疾病	1 245（11.53）	1 200（96.39）	3.84 ± 2.21
胃部疾病或消化系统疾病	3 403（31.51）	3 069（90.19）	2.85 ± 2.02
哮喘	855（7.92）	845（98.83）	4.04 ± 2.15
与记忆相关的疾病	656（6.07）	633（96.49）	4.17 ± 2.35

表 3 两病的关联规则分析及卡方检验

Table 3 Analysis of association rules and chi-square test for two chronic diseases

前项集	后项集	例数	观察发病率（支持度）（%）	期望发病率（%）	χ^2 值	P 值	观察发病率 / 期望发病率
关节炎或风湿病	高血压	2 428	22.48	20.99	39.03	<0.001	1.07
胃部疾病或消化系统疾病	关节炎或风湿病	2 056	19.04	14.03	507.53	<0.001	1.36
血脂异常	高血压	1 812	16.78	11.54	641.94	<0.001	1.45
心脏病	高血压	1 752	16.22	11.71	470.52	<0.001	1.38
胃部疾病或消化系统疾病	高血压	1 613	14.94	14.85	0.12	0.729	1.01
心脏病	关节炎或风湿病	1 428	13.22	11.06	108.62	<0.001	1.20
血脂异常	关节炎或风湿病	1 320	12.22	10.90	41.13	<0.001	1.12
慢性肺部疾患	关节炎或风湿病	1 204	11.15	8.74	159.66	<0.001	1.28
心脏病	胃部疾病或消化系统疾病	1 179	10.92	7.83	254.43	<0.001	1.39
血脂异常	心脏病	1 168	10.81	6.08	698.65	<0.001	1.78
糖尿病或血糖升高	高血压	1 144	10.59	7.23	376.91	<0.001	1.47
慢性肺部疾患	高血压	1 103	10.21	9.26	24.83	<0.001	1.10

表 4 三病组合的关联规则分析

Table 4 Analysis of association rulesfor triple disease combinations

前项集	后项集	个数	观察发病率（支持度）（%）	期望发病率（%）	观察发病率 / 期望发病率	提升度
关节炎或风湿病，胃部疾病或消化系统疾病	高血压	1011	9.40	6.60	1.42	1.42
心脏病，关节炎或风湿病	高血压	940	8.70	5.20	1.67	1.67
关节炎或风湿病，血脂异常	高血压	914	8.50	5.10	1.67	1.65
心脏病，血脂异常	高血压	895	8.30	2.90	2.86	2.89
心脏病，胃部疾病或消化系统疾病	关节炎或风湿病	798	7.40	3.50	2.11	2.12
心脏病，胃部疾病或消化系统疾病	高血压	734	6.80	3.70	1.84	1.84
糖尿病或血糖升高，血脂异常	高血压	692	6.40	1.80	3.56	3.62
血脂异常，胃部疾病或消化系统疾病	高血压	677	6.30	3.60	1.75	1.72
慢性肺部疾患	高血压	670	6.20	4.10	1.51	1.51
慢性肺部疾患	关节炎或风湿病	665	6.20	2.80	2.21	2.24
心脏病，血脂异常	关节炎或风湿病	653	6.00	2.70	2.22	2.23
血脂异常，胃部疾病或消化系统疾病	关节炎或风湿病	648	6.00	3.40	1.76	1.75
糖尿病或血糖升高，关节炎或风湿病	高血压	566	5.20	3.20	1.63	1.63
心脏病，血脂异常	胃部疾病或消化系统疾病	547	5.10	1.90	2.68	2.64

2.3 聚类分析 通过分析 14 种常见慢病的谱系图（图 1），通过距离接近程度将慢病分为 6 类，以不同颜色区分。Yule′Q 距离体现了慢病在人员中的发病信息，距离较近慢病即在被访者中共同发病率较高，发病因素多有所关联，如哮喘与慢性肺部疾患同为呼吸系统疾病，血脂异常、糖尿病或血糖升高、中风与高血压同为血管循环系统疾病，与记忆相关的疾病、情感及精神方面问题为精神心理疾病，胃部疾病或消化系统疾病与关节炎或风湿病同为代谢系统疾病。心脏病与与记忆相关的疾病、情感及精神方面问题为精神心理疾病距离接近，三者可以聚为一类，癌症等恶性肿瘤独立于其他疾病，单独为一类。



图 1 14 种慢病的系统聚类谱系图

Figure 1 Dendrogram of systematic clustering of 14 chronic diseases

被访者的聚类使用簇内平方误差和确定最佳的簇数为 4 个，聚类后每个被访者有 1 个分类，对 4 个分类的患病情况及疾病个数统计如表 5。由于 k-medoids 聚类选取 Jaccard 距离，侧重于体现同时发病信息。分类

1 包含了所有的无慢病及患有 1~2 种共病率较低的慢病的被访者，相对健康。分类 2 呼吸系统疾病、胃部疾病或消化系统疾病的发病率远高于总体，关节炎或风湿病发病率略高于总体，系统聚类中呼吸系统疾病为一类，与其他疾病相对独立。胃部疾病或消化系统疾病、关节炎或风湿病发病也独立于其他疾病，这 4 种慢病组合随机共存的患者多在分类 2 中。分类 3 中 3~5 种共病被访者占比较高，其中心脏病、血脂异常的发病率远高于总体，而高血压、中风等疾病的发病率略高于总体。常见的 3~5 种共病多为心血管疾病（高血压、糖尿病或血糖升高、血脂异常、心脏病），可以推测分类 3 为包含这几个疾病的共病组合。分类 4 中被访者均患有高血压，但是平均共病个数为 1.82 个，并且除高血压之外的慢病发病率均低于总体，可以推测分类 4 包含仅患有高血压的被访者及高血压与非心血管疾病组合的患者。

2.4 通过主成分对被访者进行 k-medoids 聚类分析 通过碎石图（图 2）确定主成分个数为 5 个，设定载荷系数 >0.36 得到各主成分不重叠的代表慢病，同一主成分的代表疾病关联性较强。第一主成分包含高血压，糖尿病或血糖升高，心脏病，血脂异常；第二主成分包含慢性肺部疾患与哮喘，第三主成分包含关节炎或风湿病，胃部疾病或消化系统疾病；第四主成分包含中风，情感及精神方面问题，与记忆相关的疾病；第五主成分包含

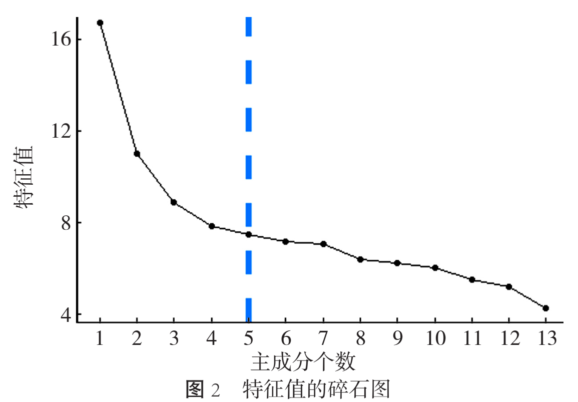


图 2 特征值的碎石图

Figure 2 Scree plot for factor analysis

癌症等恶性肿瘤，肝脏疾病，肾脏疾病。

2.5 潜在类别分析 通过最小的 BIC 指标确定最优的分类个数为 5 个，分析潜在类别的条件概率分布图（图 3）及各类别与慢病及慢病个数的交叉表（表 7）。分类 5 是为相对健康组，包含了所有的无慢病及只有 1 个慢病的被访者，大部分两病的被访者，14 个慢病的发病率均低于总体。分类 1 为代谢系统疾病组，关节炎或风湿病、胃部疾病或消化系统疾病的发病率高；分类 2 为血管循环疾病（高血压病、糖尿病或血糖升高、心脏病、血脂异常）组合；分类 3 为多病组，有 5 个以上的慢病，人数较少，多个慢病的发病率高于总体；分类 4 为呼吸

表 5 多病被访者的 k-medoids 聚类分析结果（分组被访者的发病率）

Table 5 Results of k-medoids clustering analysis of respondentswith multimorbidity

分类	例数	高血压〔例（%）〕	糖尿病或血糖升高〔例（%）〕	癌症等恶性肿瘤〔例（%）〕	慢性肺部疾患〔例（%）〕	心脏病〔例（%）〕	中风〔例（%）〕	情感及精神方面问题〔例（%）〕	关节炎或风湿病〔例（%）〕	血脂异常〔例（%）〕
分类 1	3 259	505 （15.50）	57 （1.75）	73（2.24）	0	75 （2.30）	154 （4.73）	62（1.90）	1 532（47.00）	87 （2.67）
分类 2	2 016	264 （13.10）	46 （2.28）	47（2.33）	1 161（57.60）	222（11.01）	104 （5.16）	77（3.82）	1139（56.5）	92 （4.56）
分类 3	3 861	2 656（68.80）	1 425（36.90）	124（3.21）	803（20.80）	2 197（56.82）	664（17.20）	232（6.01）	2 135（55.30）	2 259（58.50）
分类 4	1 664	1 664（100.00）	128 （7.69）	35（2.10）	156 （9.38）	191（11.48）	170（10.22）	27（1.62）	0	206（12.38）
总体	10 800	5 091（47.14）	1 656（15.33）	279（2.58）	2 121（19.64）	2 684（24.85）	1 093（10.12）	398（3.69）	4 808（44.52）	2 644（24.48）

分类	肝脏疾病〔例（%）〕	肾脏疾病〔例（%）〕	胃部疾病或消化系统疾病〔例（%）〕	哮喘〔例（%）〕	与记忆相关的疾病〔例（%）〕	无慢病〔例（%）〕	1 种慢病〔例（%）〕	2 种慢病〔例（%）〕	3 种慢病〔例（%）〕	4 种慢病〔例（%）〕
分类 1	99 （3.04）	178 （5.46）	0	44 （1.35）	80 （2.45）	1 486（45.60）	858（26.33）	688（21.11）	195 （5.98）	32 （0.98）
分类 2	172 （8.53）	242（12.00）	1 407（69.80）	333 （16.5）	81 （4.02）	0	472（23.41）	582（28.87）	443（21.97）	261（12.95）
分类 3	440（11.40）	741（19.20）	1 819（47.10）	425（11.00）	432（11.20）	0	295 （7.64）	261 （6.76）	890（23.05）	915（23.7）
分类 4	66 （3.97）	86 （5.17）	180（10.82）	55 （3.31）	63 （3.79）	0	633（38.04）	715（42.97）	300（18.03）	16 （0.96）
总体	779 （7.21）	1 245（11.53）	3 403（31.51）	855 （7.92）	656 （6.07）	1 486（13.76）	2 258（20.91）	2 246（20.96）	1 828（16.93）	1 224（11.33）

分类	5 种慢病〔例（%）〕	6 种慢病〔例（%）〕	7 种慢病〔例（%）〕	8 种慢病〔例（%）〕	9 种慢病〔例（%）〕	10 种慢病〔例（%）〕	11 种慢病〔例（%）〕	12 种慢病〔例（%）〕	13 种慢病〔例（%）〕	平均慢病种数（x ± s）
分类 1	0	0	0	0	0	0	0	0	0	0.90 ± 0.99
分类 2	194 （9.62）	50 （2.48）	10（0.50）	1（0.05）	2（0.10）	0	1（0.05）	0	0	2.67 ± 1.42
分类 3	617（15.98）	426（11.03）	247（6.40）	117（3.03）	56（1.45）	22（0.57）	8（0.21）	4（0.10）	3（0.08）	4.24 ± 1.90
分类 4	0	0	0	0	0	0	0	0	0	1.82 ± 0.75
总体	811 （7.51）	476 （4.41）	257（2.38）	118（1.09）	58（0.54）	22（0.20）	9（0.08）	4（0.04）	3（0.03）	2.57 ± 1.99

表 6 主成分分类表

Table 6 Classification table of principal component

分类	慢病	RC1	RC2	RC3	RC4	RC5
分类 1	高血压病	0.66				
	糖尿病或血糖升高	0.63				
	心脏病	0.47				
	血脂异常	0.69				
分类 2	关节炎或风湿病		0.68			
	胃部疾病或消化系统疾病		0.72			
分类 3	慢性肺部疾患			0.82		
	哮喘			0.85		
分类 4	中风				0.45	
	情感及精神方面问题				0.67	
	与记忆相关的疾病				0.74	
分类 5	癌症等恶性肿瘤					0.78
	肝脏疾病					0.50
	肾脏疾病					0.42

系统疾病组，呼吸系统疾病（慢性肺部疾患如慢性支气管炎或肺气肿、肺心病、哮喘）的患病率较高。

3 讨论

3.1 不同研究方法所得共病模式的差异及一致性

共病概念说法不一，本研究特指“Multimorbidity”，为“同一个体共存的急性病或慢病”[2]，强调共存，病因学关系上有可能相关也可能独立。本研究得到的有一致性的几种共病模式是：（i）高血压、糖尿病或血糖升高血脂异常；（ii）慢性肺部疾患、哮喘；（iii）关节炎或风湿病、胃部疾病或消化系统疾病；（iv）中风、与记忆相关的疾病。不同方法得到组合有所差异，例如心脏病在主成分分析及潜在类别分析中与模式 i 关系紧密，但在系统聚类中关联较远。恶性肿瘤在多个方法中均独立于其他疾病，情感及精神方面问题发病率较低，这两者在某些方法中与其他疾病有一些关联，但是没有明显及稳定的共病组合。肾脏疾病、肝脏疾病虽然多种分类中有所体现，但两者的病因学关系相对独立。

不同方法的原理不同。关联规则通过控制前项集及后项集的个数，得到指定数量的疾病组合，提升度判断独立性以区分共存类型，但是支持度仅体现共病观察发病率。系统聚类及主成分分析，侧重于慢病共现的关联性，结果能较为清晰的体现病因关联；k-medoids 及 LCA 侧重于通过人群反映共病组合，两者的目的均是将个体分类。k-medoids 聚类基于距离按空间位置进行分

表 7 潜在类别分析分类与慢病交叉表

Table 7 Cross table of latent classes with chronic diseases

分类	例数	高血压〔例（%）〕	糖尿病或血糖升高〔例（%）〕	癌症等恶性肿瘤〔例（%）〕	慢性肺部疾患〔例（%）〕	心脏病〔例（%）〕	中风〔例（%）〕	情感及精神方面问题〔例（%）〕	关节炎或风湿病〔例（%）〕	血脂异常〔例（%）〕
class1	1 266	501（39.6）	82（6.5）	42（3.3）	409（32.3）	400（31.6）	94（7.4）	74（5.8）	1 111（87.8）	221（17.5）
class2	2 728	2 389（87.6）	1 090（40）	73（2.7）	275（10.1）	1 287（47.2）	694（25.4）	77（2.8）	1 059（38.8）	1 623（59.5）
class3	559	446（79.8）	240（42.9）	32（5.7）	361（64.6）	497（88.9）	163（29.2）	139（24.9）	445（79.6）	437（78.2）
class4	617	304（49.2）	63（10.2）	12（1.9）	589（95.3）	259（41.9）	33（5.3）	22（3.6）	335（54.2）	111（18.00）
class5	5 629	1 451（25.8）	181（3.2）	120（2.1）	487（8.7）	241（4.3）	109（1.9）	86（1.5）	1 858（33.00）	252（4.5）
总体	10 800	5 091（47.2）	1 656（15.33）	279（2.58）	2 121（19.64）	2 684（24.85）	1 093（10.12）	398（3.69）	4 808（44.52）	2 644（24.48）

分类	肝脏疾病〔例（%）〕	肾脏疾病〔例（%）〕	胃部疾病或消化系统疾病〔例（%）〕	哮喘〔例（%）〕	与记忆相关的疾病〔例（%）〕	0 种慢病〔例（%）〕	1 种慢病〔例（%）〕	2 种慢病〔例（%）〕	3 种慢病〔例（%）〕	4 种慢病〔例（%）〕
class1	288（22.7）	346（27.3）	1 182（93.4）	29（2.3）	71（5.6）	0	0	9（0.71）	561（44.31）	402（31.75）
class2	133（4.9）	284（10.4）	575（21.1）	49（1.8）	270（9.9）	0	0	525（19.24）	896（32.84）	650（23.83）
class3	204（36.5）	328（58.7）	429（76.7）	199（35.6）	198（35.4）	0	0	0	0	0
class4	30（4.9）	60（9.7）	212（34.3）	546（88.3）	35（5.7）	0	0	51（8.27）	150（24.31）	169（27.39）
class5	124（2.2）	227（4.00）	1 005（17.9）	32（0.6）	82（1.5）	1 486（26.4）	2 258（40.11）	1 661（29.51）	221（3.93）	3（0.05）
总体	779（7.21）	1 245（11.53）	3 403（31.51）	855（7.92）	656（6.07）	1 486（13.76）	2 258（20.91）	2 246（20.8）	1 828（16.93）	1 224（11.33）

分类	5 种慢病〔例（%）〕	6 种慢病〔例（%）〕	7 种慢病〔例（%）〕	8 种慢病〔例（%）〕	9 种慢病〔例（%）〕	10 种慢病〔例（%）〕	11 种慢病〔例（%）〕	12 种慢病〔例（%）〕	13 种慢病〔例（%）〕	平均慢病种数（x ± s）
class1	224（17.69）	69（5.45）	1（0.08）	0	0	0	0	0	0	3.83 ± 0.92
class2	437（16.02）	186（6.82）	33（1.21）	1（0.04）	0	0	0	0	0	3.62 ± 1.23
class3	18（3.22）	132（23.61）	196（35.06）	117（20.93）	58（10.38）	22（3.94）	9（1.61）	4（0.72）	3（0.54）	7.37 ± 1.36
class4	132（21.39）	89（14.42）	27（4.38）	0	0	0	0	0	0	4.22 ± 1.31
class5	0	0	0	0	0	0	0	0	0	1.11 ± 0.84
总体	811（7.51）	476（4.41）	257（2.38）	118（1.09）	58（0.54）	22（0.20）	9（0.08）	4（0.04）	3（0.03）	2.57 ± 1.99

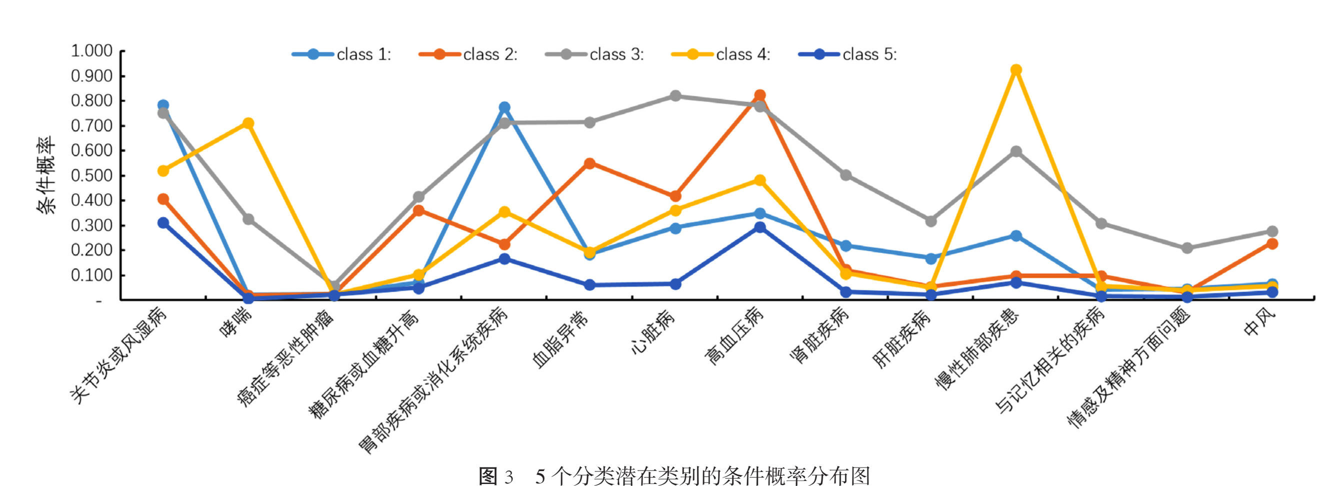


图 3 5 个分类潜在类别的条件概率分布图
Figure 3 Probability distribution graph of five latent classes

类，分类结果与所选距离算法有关；LCA 相对客观，是注重于个体在类别显变量上的不同反应模式对个体进行分类、识别群体异质性的一种统计方法[16]。两种方法通过人群发病率得到共病组合。

各方法所得结果的差异性可以解释。例如关联规则所得的两病组合：高血压与关节炎或风湿病，观察发病率最高，但系统聚类、被访者 k-medoids 聚类、主成分分析中均未被分为一类。这是由于高血压与关节炎或风湿病在病因学关系上独立，仅是随机共存的概率较高。关联规则中以高血压为后项集得到的另外几个两病组合（血脂异常与高血压、心脏病与高血压、糖尿病或血糖升高与高血压），发病率与期望发病率的比值 >1，病因学关系上并不独立，致病因素相关，在其他几种方法中也被分为一类。胃部疾病或消化系统疾病、关节炎或风湿病这一共病组合，发病率与期望发病率的比值 >1，这一共病组合在其他方法中也出现。关联规则方法简单直接但单纯关注发病率会忽略一些关联紧密但发病率较低的共病模式。例如慢性肺部疾患如慢性支气管炎或肺气肿、肺心病与哮喘，由于哮喘的发病率较低，共病率在人群中不高，但是其他几种方法均能体现此组合。

3.2 本研究与同类研究的差异及一致性 本研究结论与国内同类研究有差异也有一致。刘贝贝等[8]研究选取了同样的数据及方法，但删除了数据缺失的被访者，年龄的选取范围为 45 岁以上，得到发病率明显较低。本研究保留缺失数据，得到结论更接近 CHARLS 被访者总体及国内 60 岁及以上人群特征。数据差异但结论一致，与高血压组合的疾病较多，血管循环系统两病组合提升度较高。黎艳娜等[17]的研究也采用了 2018 年的 CHARLS 数据（数据版本不同，导致数据量有轻微差别），通过共现网络图得到了两病及三病组合，本研究数据及结论与其一致。孙至佳等[7]的研究基于中国慢病前瞻性研究（China Kadoorie Biobank，CKB）数据，以 Yule′s Q 距离进行系统聚类，得到基线调查数据中，糖尿病、冠心病、中风及高血压聚为一类的结论，本研究结果与之一致。但是第二次重复调查中，结论有所区别，本研究恶性肿瘤与其他慢病的关系也与其不同。

同类研究结果的不一致在于疾病有所不同，也由于调查人员与被访者的认知及对疾病的理解不同，访谈数据存在一定的误差。本研究中多个方法均显示恶性肿瘤与其他疾病并没有明显的共病特性，而孙至佳等[7]的研究中恶性肿瘤与其他疾病的关联更紧密，这个差异可能是由于恶性肿瘤的定义不同以及被访者人群不同。

3.3 本研究的独特与不足 本研究的独特之处在于数据总量在国内同类研究中相对较大，使用一系列方法进行多病分类，并尝试新算法。除了系统聚类采用谱系图与临床结合的方法判别分类个数，被访者 k-medoids 聚类采用了体现数据特性的方法进行个数判断，主成分分析采用碎石图确定个数，LCA 采用遍指标法历法选择分类数，减少了主观判断。国内共病模式相关分析方法少，以慢病距离算法为例，国外多采用 Jaccard 以及 Yule′s Q 距离，国内较少采用。后两种算法更能体现 01 数据包含的信息。

本研究尝试通过对比结果来说明方法的适用性，得到的一致共病组合病因关联明显。主成分和系统聚类对慢病分类，而 k-medoids 及 LCA 对人群分类，可以进一步关联患者特征深入分析。聚类分析依赖于距离的计算方法，而潜在类分析用的潜变量来解释外显指标之间的关联，被潜变量解释后的外显指标间实现局部独立性，更加客观，保证了多病模式分类的稳定性。多病研究需要根据研究目的进行选择模式分析方法，或者使用多种方法得到具有一致性的稳定分类。

本研究的不足之处在于为了保证数据客观性及样本数量，通过众数填充多病缺失数据，会影响发病率的准确度，导致发病率低于总体。保留了缺失的被访者人口

特征学数据并没有进行处理，虽然分类到被访者个人，但未结合人群的各种特征进一步研究，未分析危险因素及预后，内容不够贴近临床。

作者贡献：王留义提出研究选题方向和总体研究目标，对研究进行可行性分析，对文章整体负责，监督管理；潘晔负责数据清洗、处理和统计学分析、R 语言编程、绘制图表；潘晔、刘志辉、胡倩倩负责结果分析与解释，撰写论文初稿；潘晔、刘志辉负责论文的修订。

本文无利益冲突。

参考文献

[1] World Health Organization.The World Health Report 2008——primary health cared （now more than ever）[M]. New York：World Health Report，2008：148.

[2] 常峰，张冉，路云，等. 共病概念体系研究分析与设计[J]. 中国全科医学，2018，21（3）：256–260. DOI：10.3969/j.issn.1007–9572.2017.00.219.

[3] WU F，GUO Y F，KOWAL P，et al. Prevalence of major chronic conditions among older Chinese adults：the study on global AGEing and adult health（SAGE）wave 1[J]. PLoS One，2013，8（9）：e74176. DOI：10.1371/journal.pone.0074176.

[4] CORNELL J E，PUGH J A，WILLIAMS J W，et al. Multimorbidity clusters：clustering binary data from multimorbidity clusters：clustering binary data from a large administrative medical database [J]. Appl Multivar Res，2009，12（3）：163. DOI：10.22329/amr.v12i3.658.

[5] SCHÄFER I，VON LEITNER E C，SCHÖN G，et al. Multimorbidity patterns in the elderly：a new approach of disease clustering identifies complex interrelations between chronic conditions [J]. PLoS One，2010，5（12）：e15941. DOI：10.1371/journal.pone.0015941.

[6] GARCÍA–OLMOS L，SALVADOR C H，ALBERQUILLA Á，et al. Comorbidity patterns in patients with chronic diseases in general practice[J]. PLoS One，2012，7（2）：e32141. DOI：10.1371/journal.pone.0032141.

[7] 孙至佳，樊俊宁，余灿清，等. 中国 10 个地区成年人共病流行特征分析[J]. 中华流行病学杂志，2021，42（5）：755–762. DOI：10.3760/cma.j.cn112338–20200305–00259.

[8] 刘贝贝，田庆丰，郭金玲. 我国中老年人群慢性病患病现状及共病模式分析[J]. 医学与社会，2022，35（8）：58–61，66.

[9] OLAYA B，MONETA M V，CABALLERO F F，et al. Latent class analysis of multimorbidity patterns and associated outcomes in Spanish older adults：a prospective cohort study [J]. BMC Geriatr，2017，17（1）：186. DOI：10.1186/s12877–017–0586–1.

[10] ZHAO Y H，STRAUSS J，YANG G H，et al.（2013）. China Health and Retirement Longitudinal Study：2011–2012 National Baseline User's Guide，National School of Development，Peking University.

[11] ZHAO Y H，STRAUSS J，CHEN X X，et al.（2020）. China Health and Retirement Longitudinal Study Wave 4 User's Guide，National School of Development，Peking University.

[12] ZHAO Y，HU Y，SMITH J P，et al. Cohort profile：the China health and retirement longitudinal study （CHARLS）[J]. Int J Epidemiol，2014，43（1）：61–68. DOI：10.1093/ije/dys203.

[13] LLOYD B P，KENNEDY C H，YODER P J. Quantifying contingent relations from direct observation data：transitional probability comparisons versus yule's Q[J]. J Appl Behav Anal，2013，46(2)：479–497. DOI：10.1002/jaba.45.

[14] A note on the triangle inequality for the Jaccard distance [J]. Pattern Recognit Lett，2019，120：36–38. DOI：10.1016/j.patrec.2018.12.007.

[15] 杨涵墨. 中国人口老龄化新趋势及老年人口新特征[J]. 人口研究，2022，46（5）：104–116. DOI：10.3969/j.issn.1000–4149.2019.01.006.

[16] 温忠麟，谢晋艳，王惠惠. 潜在类别模型的原理、步骤及程序[J]. 华东师范大学学报：教育科学版，2023，41（1）：1–15. DOI：10.16382/j.cnki.1000–5560.2023.01.001.

[17] 黎艳娜，王艺桥. 我国老年人慢性病共病现状及模式研究[J]. 中国全科医学，2021，24（31）：3955–3962，3978. DOI：10.12114/j.issn.1007–9572.2021.00.295.

（收稿日期：2023–02–24；修回日期：2023–05–10）

（本文编辑：贾萌萌）